



## Science Arts & Métiers (SAM)

is an open access repository that collects the work of Arts et Métiers ParisTech researchers and makes it freely available over the web where possible.

This is an author-deposited version published in: <http://sam.ensam.eu>  
Handle ID: <http://hdl.handle.net/10985/9698>

### To cite this version :

Adil JARRAH, Jean Marie NIANGA, Alain IOST, Gildas GUILLEMOT, Denis NAJJAR - On the detection of corrosion pit interactions using two-dimensional spectral analysis - Corrosion Science - Vol. 52, n°2, p.303-313 - 2010

Any correspondence concerning this service should be sent to the repository  
Administrator : [archiveouverte@ensam.eu](mailto:archiveouverte@ensam.eu)

---

# On the detection of corrosion pit interactions using two-dimensional spectral analysis

Adil Jarrah <sup>a,b,\*</sup>, Jean-Marie Nianga <sup>a</sup>, Alain Iost <sup>b</sup>, Gildas Guillemot <sup>b</sup>, Denis Najjar <sup>b</sup>

<sup>a</sup> Equipe Mécanique des Structures, HEI, 13 rue de Toul, 59046 Lille Cedex, France

<sup>b</sup> Arts et Métiers ParisTech, CNRS, LMPGM, Equipe Caractérisation et Propriétés de la Pêrisurface, 8 Boulevard Louis XIV, 59046 Lille Cedex, France

---

## A B S T R A C T

A statistical methodology for detecting pits interactions based on a two-dimensional spectral analysis is presented. This method can be used as a tool for the exploratory analysis of spatial point patterns and can be advanced as an alternative of classical methods based on distance. One of the major advantages of the spectral analysis approach over the use of classical methods is its ability to reveal more details about the spatial structure like the scale for which pits corrosion can be considered as independent. Furthermore, directional components of pattern can be investigated.

The method is validated in a first time using numerical simulations on random, regular and aggregated structures. The density of pits, used in the numerical simulations, corresponds to that assessed from a corroded aluminium sheet. In a second time, this method is applied to verify the independence of the corrosion pits observed on the aforementioned aluminium sheet before applying the Gumbel theory to determine the maximum pit depth. Indeed, the property of independence is a prerequisite of the Gumbel theory which is one of the most frequently used in the field of safety and reliability.

---

---

### Keywords:

C. Pitting corrosion  
Spatial patterns  
Spectral analysis  
Pits interaction  
Gumbel's method

---

## 1. Introduction

Pitting corrosion is one of the most prevalent forms of localized corrosion. This stochastic physical phenomenon can be analyzed using probabilistic models and statistical approaches in order to predict the maximum pit depth value from measurements collected on smaller analyzed surfaces  $s$ . The extreme values theory is one of the most used for this aim [1]. However, its theoretical conditions for a rigorous application are rarely respected or not easy to check, notably homogeneity and independence. The homogeneity states that the probability distribution of the random variable representing the depth is the same for all pits, while the independence supposes that there is no interaction between random variables, meaning that the evolution of a pit doesn't influence that of the others.

The independence quantification of pits seems to be a questionable assumption with respect to the extreme value theory because of the complexity of the corrosion phenomenon. Indeed, several methods were developed to characterize existing interac-

tions among pits based on spatial point process. Lopez De La Cruz et al. [2] applied the inter-event distances estimator method, Budiansky et al. [3] and Cawley et al. [4] analyse pits interaction by using the Ripley's method. Namely, all of these methods are based on distance measure and offer a statistical treatment allowing to distinguish between different types of spatial data: the aggregated, the regular and the random spatial structures.

In this paper, an alternative method based on a two-dimensional spectral analysis is presented in order to investigate the structure of a spatial point pattern. Since the seminal work of Bartlett [5], the spectral analysis has been developed with the advent of powerful computers, and it was more widely used in particular in the domains of plant ecology and of forest management [6,7]. This method provides a detailed description of a structure and has the advantage to detect the presence of directional components.

In Section 2, a theoretical background is presented. The three common types of spatial data are reminded, then classical methods (quadrat counts and Ripley's method) are presented, and at last, the spectral analysis is detailed. The relevancy of this method is showed using numerical simulations which is presented in Section 3. Afterwards, the experimental procedures are detailed with a discussion about the different results in Section 4. Finally, a conclusion is presented in Section 5.

---

\* Corresponding author. Address: Equipe Mécanique des Structures, HEI, 13 rue de Toul, 59046 Lille Cedex, France. Tel.: +33 320622233; fax: +33 32062257.

E-mail addresses: [adil.jarrah@yahoo.fr](mailto:adil.jarrah@yahoo.fr) (A. Jarrah), [jean-marie.nianga@hei.fr](mailto:jean-marie.nianga@hei.fr) (J.-M. Nianga), [alain.iost@lille.ensam.fr](mailto:alain.iost@lille.ensam.fr) (A. Iost), [gildas.guillemot@lille.ensam.fr](mailto:gildas.guillemot@lille.ensam.fr) (G. Guillemot), [denis.najjar@lille.ensam.fr](mailto:denis.najjar@lille.ensam.fr) (D. Najjar).

## 2. Theoretical background

### 2.1. Types of spatial data

Generally, three types of spatial structures can be distinguished: random, regular and aggregated. An example of these structures will be presented later.

A random spatial structure is known as the complete spatial randomness (CSR). It serves as the null hypothesis and asserts that:

(a) The number of events  $k$  (points) belonging to a region  $A$  with area  $|A|$  follows a Poisson distribution with mean  $\lambda|A|$ , where  $\lambda$  represents the density which is the mean number of points per unit area. The probability of having  $k$  points into the region  $A$  is

$$P_k(A) = \exp(-\lambda |A|) \frac{(\lambda |A|)^k}{k!} \quad k \in \mathbb{N} \quad (1)$$

(b) Given  $N$  events within  $A$ , the locations of these events are independent and uniformly distributed over  $A$ .

The stochastic process which generates the CSR structures is the homogeneous Poisson point process denoted by HPP. The main properties of this process are stationarity and isotropy. A point process  $X = \{x_i\}_{i=1}^n (x_i \in \mathbb{R}^2, i = 1, \dots, n)$  is stationary if the translated process  $X_z = \{x_i + z\}_{i=1}^n$  has the same distribution for all  $z \in \mathbb{R}^2$ . A process is isotropic if its distribution is invariant with respect to rotations. A process which possesses these properties is said motion-invariant [8]. The HPP is an example of such processes.

In the context of pitting corrosion, if the null hypothesis is rejected, this means that the independence assumption is violated. So, the alternative hypothesis is privileged and it is assumed that the process which generates pitting corrosion is either regular or clustered. The first process supposes that every pit is located at least at for a distance denoted later by  $\delta$  and the second is characterized by the existence of many sites of clustered pits. Both of these processes violate the hypothesis of independence.

### 2.2. Classical methods

#### 2.2.1. Ripley's method

Among the methods based on distance, the Ripley's method can characterize the spatial structure of a point pattern. This last method aims to assess the probability to observe a number of points belonging to a circle of radius  $r$ . Under the hypothesis of homogeneity, the local density in every circle is equal to the total density of the studied region  $\Omega$ . The Ripley's function measures the fluctuations by drawing circles of increasing radius  $r$ . It is defined as:

$$K(r) = \frac{E(N_n(r))}{\lambda} \quad (2)$$

where  $N_n(r)$  is the random variable corresponding to the number of events into the circle of radius  $r$  and  $E(\cdot)$  is the operator of expectation.

Therefore, for a CSR structure, we have  $K(r) = \pi r^2$ , for an aggregated structure  $K(r) > \pi r^2$  and for a regular structure  $K(r) < \pi r^2$ . In practice the function  $K(r)$  is unknown. However, a naive estimator of this function can be obtained by:

$$\hat{K}(r) = \frac{1}{\hat{\lambda}} \frac{1}{N} \sum_{i=1}^N \sum_{j=1, j \neq i}^N 1(d_{ij} \leq r) \quad (3)$$

where  $\hat{\lambda} = \frac{N}{|\Omega|}$  is the estimator of the density,  $N$  the total number of points,  $|\Omega|$  the area of the studied region,  $d_{ij}$  the distance between events  $i$  and  $j$  and  $1(d_{ij} \leq r)$  is the indicator function defined as:

$$1(d_{ij} \leq r) = \begin{cases} 1 & \text{if } d_{ij} \leq r \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

This estimator (3) is biased because of the edge effects. These effects arise because the theoretical distributions for most spatial point statistics assume an unbounded area. Thus, corrections are required because the points near to the edges have fewer neighbours. As an example, Ripley [9] proposed the following correction:

$$\hat{K}(r) = \frac{1}{\hat{\lambda}} \frac{1}{N} \sum_{i=1}^N \sum_{j=1, j \neq i}^N \frac{1(d_{ij} \leq r)}{w_{ij}} \quad (5)$$

where  $w_{ij}$  is the proportion of the circumference of the circle centred at event  $i$  of radius  $d_{ij}$  which is contained within  $A$ .

In practice, the upper and the lower confidence bounds (confidence envelopes) are calculated using Monte-Carlo simulations. Namely, several CSR structures are generated for which the Ripley's functions are calculated for the various values of  $r$ . For a given risk  $\alpha$  and a radius  $r$ , the upper confidence bound corresponds to the  $1 - \frac{\alpha}{2}$  quantile and the lower confidence bound to the  $\frac{\alpha}{2}$  quantile.

A CSR structure is concluded from the observed pattern if the estimator  $\hat{K}(r)$  remains inside the envelope, an aggregated structure if  $\hat{K}(r)$  falls above the upper confidence bound and a regular structure if  $\hat{K}(r)$  falls below the lower confidence bound.

The main problem of the Ripley's method is the bias related to the edge effects and to the increasing of distance  $r$ . Thus, corrections must be taken into account.

#### 2.2.2. Quadrat counts method

This method supposes that the studied region  $\Omega$  is a rectangle (or a square). It consists in subdividing the studied region  $\Omega$ , containing  $n_{\text{total}}$  points, into  $m$  cells with the same area (quadrats). The number  $n_i$  of points belonging to the cell  $i$  ( $i = 1, \dots, m$ ) is then counted. So, under CSR hypothesis, the expected value of  $n_i$  for  $i = 1, \dots, m$  is given by:

$$\bar{n} = \frac{n_{\text{total}}}{m} \quad (6)$$

Therefore, the variable  $D$  defined as:

$$D = \sum_{i=1}^m \frac{(n_i - \bar{n})^2}{\bar{n}} \quad (7)$$

follows the Chi-square distribution with  $m-1$  degrees of freedom.

The main problem encountered in this method is the choice of the size of quadrat. Namely, a CSR hypothesis can be rejected or accepted for two different choices of  $m$ . In order to reduce the drawbacks of this method, Lopez De La Cruz et al. [10] coupled it with the non homogeneous Poisson process.

Finally, it should be noticed that the methods mentioned previously can lead to contradictory results like pit pattern studied in [2]. For this reason, Cox [11] recommended to analyse the structures by using various methods. According to this scope, spectral analysis is then performed and applied to pitting corrosion.

#### 2.2.3. Spectral analysis of spatial point processes

As mentioned previously, spectral analysis of spatial data can be presented as an alternative to the classical methods based on distance. In order to perform this method, a mathematical background is proposed which is based on the work of Mugglestone et al. [12]. Diggle's notation and terminology are used here [13].

In order to distinguish between different spatial point patterns, it is common to compare their first and second order properties.

**2.2.3.1. First and second order properties. First order property:** The first order property called also the intensity function is defined as the expected number of points per unit area which is given by:

$$\lambda_X(a) = \lim_{|da| \rightarrow 0} \left\{ \frac{E[N_X(da)]}{|da|} \right\}, \quad a = (a_1, a_2) \in \mathbb{R}^2 \quad (8)$$

where  $da$  is the elementary surface containing the point  $a$ .  $N_X(da)$  represents the number of events generated by a spatial process  $X$  in  $da$  and  $|da|$  represents the area of  $da$ . In the case of stationary process, the intensity function  $\lambda_X(a)$  is constant and denoted later by  $\lambda_X$ .

**Second order property:** The second order property contains information about the dependence between events in two regions. Diggle defined it as:

$$\lambda_{XX}(a, b) = \lim_{|da||db| \rightarrow 0} \left\{ \frac{E[N_X(da)N_X(db)]}{|da||db|} \right\}, \quad a \neq b; a, b \in \mathbb{R}^2 \quad (9)$$

For theoretical description of second order property, the covariance density function is more useful. For  $a, b \in \mathbb{R}^2$ , it is defined as:

$$\gamma_{XX}(a, b) = \lim_{|da||db| \rightarrow 0} \left\{ \frac{E[\{N_X(da) - \lambda_X(a)|da|\}\{N_X(db) - \lambda_X(b)|db|\}]}{|da||db|} \right\}, \quad a \neq b \quad (10)$$

which is equivalent to:

$$\gamma_{XX}(a, b) = \lambda_{XX}(a, b) - \lambda_X(a)\lambda_X(b), \quad a \neq b \quad (11)$$

The definition of the covariance density function in (9) is available when  $a \neq b$ . The extension to the case  $a = b$  assumes that the process is orderly which means that only one event can occupy a particular point in space [12]. This case is taken into account by Bartlett [5] who has introduced the complete covariance density function:

$$\kappa_{XX}(a, b) = \lambda_X(a)\delta(a - b) + \gamma_{XX}(a, b), \quad a, b \in \mathbb{R}^2 \quad (12)$$

where  $\delta$  is the two-dimensional Dirac function for which:

$$\int_{\mathbb{R}^2} f(x)\delta(x - x')dx = f(x') \quad (13)$$

for all continuous function  $f$ . A full proof of the formula (12) is detailed on [14].

### 2.2.3.2. Practice of spectral analysis.

#### (a) Definition and calculation of the periodogram.

The spectral density function of a stationary process is defined as the Fourier Transform of the complete covariance density function  $\kappa_{XX}$  [12]:

$$f_{XX}(w_1, w_2) = \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} \kappa_{XX}(c_1, c_2) \exp\{-i(w_1 c_1 + w_2 c_2)\} dc_1 dc_2 \quad (14)$$

More details about the expressions of  $\kappa_{XX}$  and  $\gamma_{XX}$  under the hypothesis of stationarity are given on [13].

Let  $\Omega = [0, L_x] \times [0, L_y]$  be a rectangular region containing  $N$  events from a spatial process  $X$ . Let  $\{(x_j, y_j), j = 1, \dots, N\}$  be the location of events. Because of the complexity of spectral density function, Bartlett [5] introduced an estimation named **Periodogram** which is based on the discrete Fourier transform (DFT). Its expression is given by [12]:

$$\hat{f}_{XX}(w_p, w_q) = F_X(p, q) \overline{F_X(p, q)} = \{A_X(p, q)\}^2 + \{B_X(p, q)\}^2 \quad (15)$$

where:

$$\begin{aligned} F_X(p, q) &= \sum_{j=1}^N \exp\left(-2i\pi\left(p \frac{x_j}{L_x} + q \frac{y_j}{L_y}\right)\right) \\ &= A_X(p, q) + iB_X(p, q), \quad (p, q) \in \mathbb{Z}^2 \end{aligned} \quad (16)$$

In order to simplify these expressions, it is common to rescale the observed pattern to the unit area. Formula (16) becomes:

$$F_X(p, q) = \sum_{j=1}^N \exp\left(-2i\pi\left(p x'_j + q y'_j\right)\right), \quad (x'_j, y'_j)_{1 \leq j \leq N} \in [0, 1]^2 \quad (17)$$

The main property of the periodogram is symmetry, which is expressed by:

$$\hat{f}_{XX}(w_{-p}, w_{-q}) = \hat{f}_{XX}(w_p, w_q) \quad (18)$$

Thus, the periodogram values are calculated over both negative and positive integers for one of the frequency, for example  $q$ , and positive integers only for the other frequency  $p$ . Mugglestone et al. [12] propose the values  $p = 0, \dots, 16$  and  $q = -16, \dots, 15$  that cover a reasonable range of frequency. Otherwise, the choices of  $p$  and  $q$  are left to the user. In this work, we choose the range of values  $p = 0, \dots, p_{\max}$  and  $q = -q_{\max}, \dots, q_{\max} - 1$  where  $p_{\max} = q_{\max} = 64$ . The interest of using  $p_{\max}$  and  $q_{\max}$  values is to allow the exploration of smaller scales patterns.

#### (b) Interpretation of the periodogram.

The values of the periodogram represent the relative contribution of every frequency  $w$  to the total spatial variance of the process. For a CSR structure, there is no value which dominates the periodogram. For a regular (respectively aggregated) structure, we observe small (respectively high) values of the periodogram for small values of  $w$ .

For a better interpretation of the periodogram, Renshaw et al. [15,16] propose the use of the polar representation described by the  $R$ -Spectrum which investigates the scales of the studied pattern and the  $\theta$ -Spectrum which inspects the presence of directional features.

#### (c) $R$ -Spectrum and $\theta$ -Spectrum.

The  $R$ -Spectrum denoted by  $\hat{f}_R(r)$  is defined as:

$$\hat{f}_R(r) = \frac{1}{n_r} \sum_{r'} \sum_{\theta} \hat{f}_{XX}(r', \theta), \quad 1 \leq r \leq \sqrt{p_{\max}^2 + q_{\max}^2} \quad (19)$$

$\hat{f}_R(r)$  averages periodogram values for all the  $n_r$  ordinates  $\hat{f}_{XX}(r', \theta)$  for which  $r-1 < r' \leq r$ . In polar representation,  $\hat{f}_{XX}(r', \theta)$  corresponds to  $\hat{f}_{XX}(p, q)$  where the pair of values  $(p, q)$  is defined as:

$$r' = \sqrt{p^2 + q^2} \quad \text{and} \quad \theta = \arctan\left(\frac{q}{p}\right)$$

where  $r'$  represents the wave number and  $\theta$  the angle  $(-90^\circ < \theta < 90^\circ)$ .

For example, to calculate  $\hat{f}_R(6)$ , we take average of all the  $n_r$  periodogram values for which  $5 < \sqrt{p^2 + q^2} \leq 6$ . The choice of a step of  $r$  equal to 1 is recommended in practice [12].

Similarly, the  $\theta$ -Spectrum, denoted by  $\hat{f}_\Theta(\theta)$ , is constructed by averaging periodogram values for which  $\theta - 5^\circ < \theta' \leq \theta + 5^\circ$ . Step equal to  $10^\circ$  of  $\theta$  is usually used [12]:

$$\hat{f}_\Theta(\theta) = \frac{1}{n_\theta} \sum_r \sum_{\theta'} \hat{f}_{XX}(r, \theta'), \quad \theta = -90^\circ, \dots, 90^\circ \quad (20)$$

Finally, it should be noticed that the ordinate at the origin  $\hat{f}_{XX}(0, 0)$  is excluded from the averaging procedure due to its large value equals to  $N^2$  which dominates the shape of the periodogram [12].

#### (d) Probability laws of $R$ -Spectrum and $\theta$ -Spectrum.

Under CSR hypothesis, the intensity function  $\lambda_X$  is constant. In practice, it is replaced by the observed intensity  $\frac{N}{|\Omega|}$ . In order to

simplify, the pattern is rescaled to the unit square. Mugglestone et al. [12,17] assert that:

$$\frac{2\hat{f}_{XX}(w_p, w_q)}{N} \sim \chi_2^2, \quad (w_p, w_q) \neq (0, 0) \quad (21)$$

and

$$\frac{2\{\hat{f}_{XX}(0, 0) - N\}}{N} \sim \chi_1^2 \quad (22)$$

where  $\chi_n^2$  represents Chi-square distribution with  $n$  degrees of freedom.

For the  $R$ -Spectrum, we obtain:

$$\frac{\hat{f}_R(r)}{N} \sim \frac{1}{2n_r} \chi_{2n_r}^2, \quad 1 \leq r \leq \sqrt{p_{max}^2 + q_{max}^2} \quad (23)$$

with

$$E\left[\frac{\hat{f}_R(r)}{N}\right] = 1 \quad (24)$$

and for the  $\theta$ -Spectrum

$$\frac{\hat{f}_\Theta(\theta)}{N} \sim \frac{1}{2n_\theta} \chi_{2n_\theta}^2, \quad \theta \in [-90^\circ, 90^\circ] \quad (25)$$

with

$$E\left[\frac{\hat{f}_\Theta(\theta)}{N}\right] = 1 \quad (26)$$

(e) Discriminating criterion of spatial structures.

Under CSR hypothesis  $\frac{\hat{f}_R(r)}{N}$  remains inside the confidence band  $\frac{1}{2n_r} \left[ \chi_{2n_r, \frac{\alpha}{2}}^2, \chi_{2n_r, 1-\frac{\alpha}{2}}^2 \right]$  where  $\alpha$  is the risk of the test.

For small values of  $r$ , an aggregated structure is detected if  $\frac{\hat{f}_R(r)}{N}$  is above the upper confidence bound, whereas a regular structure is detected if  $\frac{\hat{f}_R(r)}{N}$  falls below the lower confidence bound.

The  $R$ -Spectrum reveals the layout of events within the studied region, while  $\theta$ -Spectrum detects directional features. The process is isotropic if  $\frac{\hat{f}_\Theta(\theta)}{N}$  remains inside the confidence bound  $\frac{1}{2n_\theta} \left[ \chi_{2n_\theta, \frac{\alpha}{2}}^2, \chi_{2n_\theta, 1-\frac{\alpha}{2}}^2 \right]$ .

### 3. Numerical simulations

In order to test the relevance of the spectral analysis method and its capacity to distinguish between different patterns of spatial data, three point processes referring to random (CSR), aggregated and regular structures are simulated. Results are compared with those given by Ripley's method for which the confidence envelope is calculated by Monte-Carlo method. It should be mentioned that all of the numerical simulations and the calculus of the different functions are implemented on Matlab software.

#### 3.1. Density of simulated structures

The density in terms of number of events per unit area used for simulations of the different structures was determined from a pitting corrosion experiment performed on an aluminium sheet of 150 cm<sup>2</sup> area. More details will be provided about this corrosion experiment in part 4.

Fig. 1 shows this corroded sheet. The calculation of the total number of pits contained within 5 squares of 4 cm<sup>2</sup> was performed visually and without any cleaning process. This calculation enables to estimate a pit density equal to 12.4 pits per cm<sup>2</sup>. To reduce the

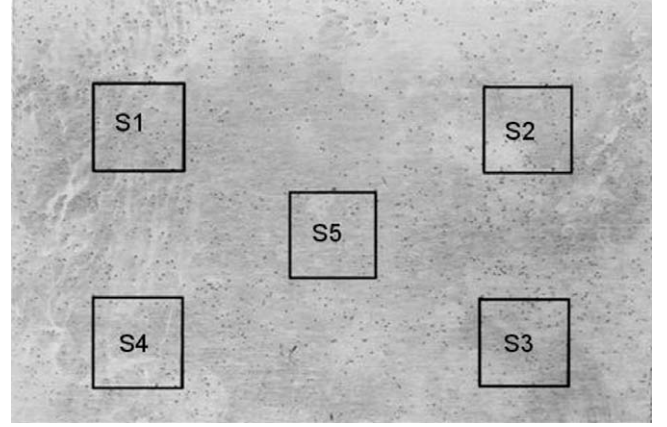


Fig. 1. Aluminium sheet  $S$  of 150 cm<sup>2</sup> area showing pitting corrosion with the selected mask.

time calculation, 200 pits were simulated in an area of 16 cm<sup>2</sup>. It should be noticed that the model consists in locating each pit with the coordinates of its mass centre called centroid hereafter.

#### 3.2. CSR structure

As said previously, the homogeneous Poisson process denoted by HPP generates the CSR structures. A realization of this process gives points distributed uniformly across the plane independently of one another [8]. Fig. 2 shows an example of 200 points (events) generated from HPP within the square  $[0,4] \times [0,4]$ . This figure is generated by simulating the HPP process via a function implemented on Matlab software. The periodogram is illustrated by Fig. 3 and the calculation of the  $R$ -Spectrum and  $\theta$ -Spectrum are summarized in Fig. 4.

The periodogram of a CSR structure has the property that none of its values dominates the others which is illustrated by Fig. 3. The confidence interval shown in Fig. 4(L) confirms the presence of CSR structure generated by the HPP since the  $R$ -Spectrum values fall within the confidence envelope. Therefore, as mentioned previously, the HPP is isotropic. This property is highlighted in Fig. 4(R) since the  $\theta$ -Spectrum remains inside the confidence envelope.

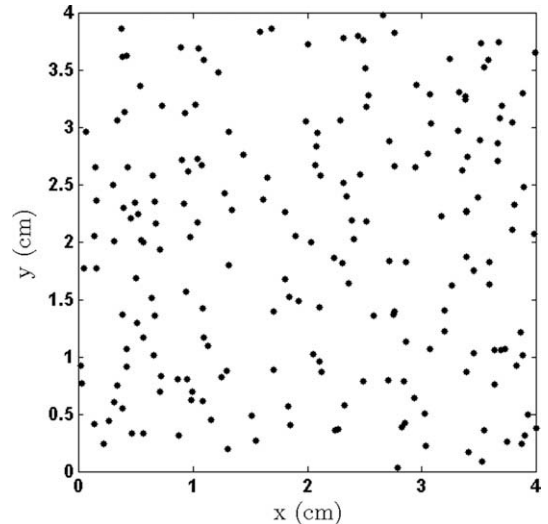


Fig. 2. Example of a realization of HPP.



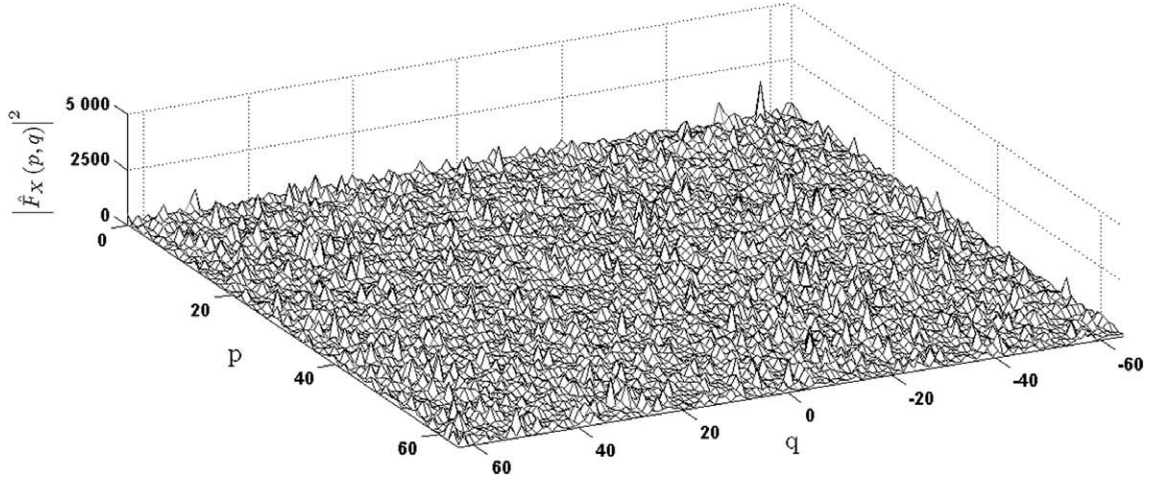


Fig. 3. Periodogram of HPP.

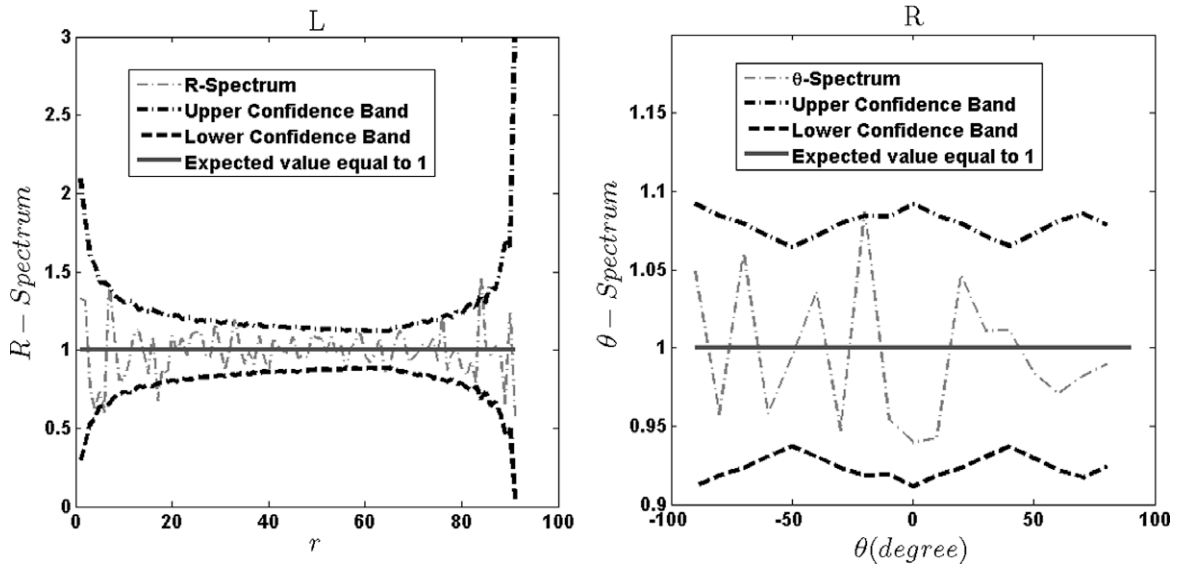


Fig. 4. 90% Confidence envelope of HPP. (L) R-Spectrum. (R)  $\theta$ -Spectrum.

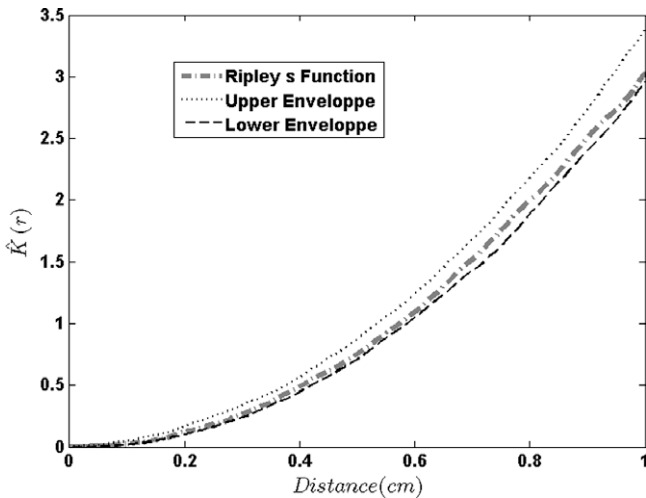


Fig. 5. 90% Confidence envelope of Ripley's method of HPP.

So as to compare spectral analysis with methods based on distance, Ripley's method is performed and the result is shown in Fig. 5.

The envelopes of Ripley's method in Fig. 5 are calculated via Monte-Carlo simulation performed 200 times. The result confirms the CSR structure already detected by spectral analysis.

### 3.3. Aggregated structure

Many stochastic models can generate cluster processes. In this work, a *Poisson Cluster Process* is used. One form of this class of processes is Modified Thomas cluster process denoted by MTCP [12]. It consists in generating the centroid of clusters called "parents" from an homogeneous Poisson process with intensity  $\lambda_p$ . Each "parent" gives rise to a number of "offsprings" which the locations are independently and identically distributed according to a bivariate probability distribution function denoted by  $h$  and their numbers follow a Poisson distribution of mean  $\mu$ .

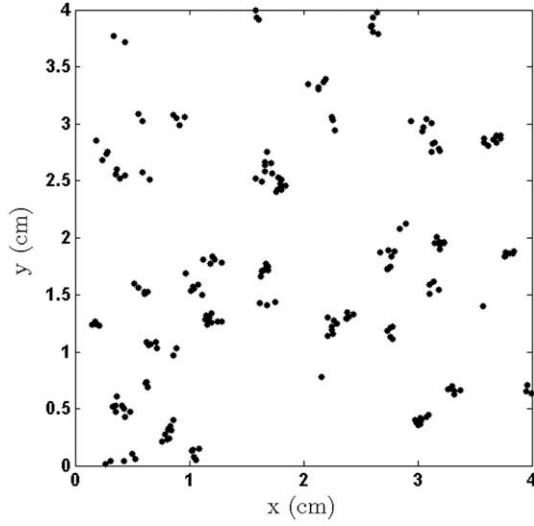


Fig. 6. Example of a realization of MTCP.

If  $h$  is symmetric, then the process is isotropic [12]. The function  $h$  is usually taken as a bivariate Gaussian distribution function for which the locations of the “offsprings” are centred on the “parent” with a standard deviation equals to  $\sigma$ :

$$h(x,y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) \quad (27)$$

where  $(x,y)$  are the coordinates of the “offsprings” relatively to the parents.

Only “offsprings” are retained in the final number and the number of events generated by MTCP is random. Then, in order to obtain a density of points close to the estimated one which is equal to 12.4 pits per unit area, we choose  $\lambda_p = 50$ ,  $\mu = 4$  and  $\sigma = 0.04$  leading to a final number of points equal to 204.

Figs. 6–8 show respectively the realization, the periodogram and the spectrums of the generated MTCP.

The periodogram of an aggregated structure is characterized by high values for small values of frequencies as illustrated in Fig. 7. About the  $R$ -Spectrum presented in Fig. 8(L), it is expected to observe high values, which fall above the upper confidence bound, for small values of  $r$ . The figure highlights this phenomenon and

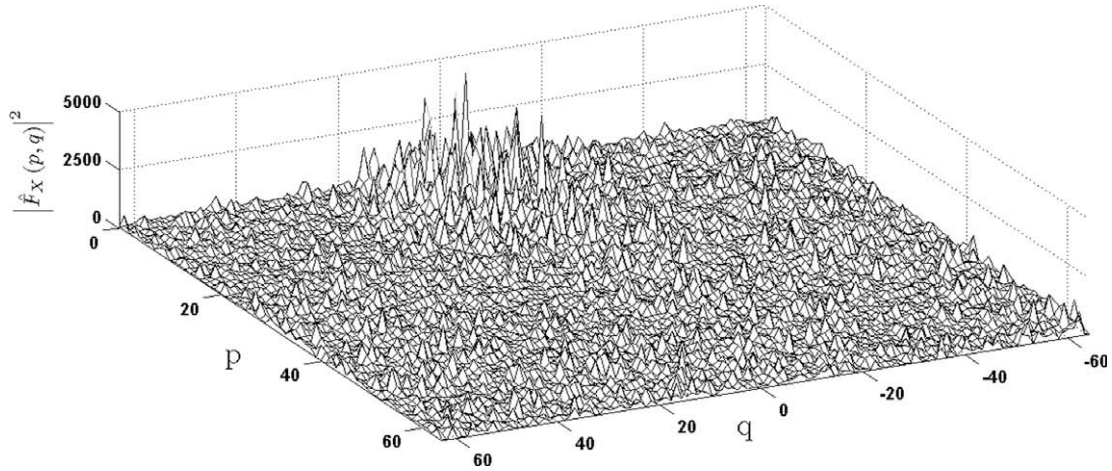


Fig. 7. Periodogram of MTCP.

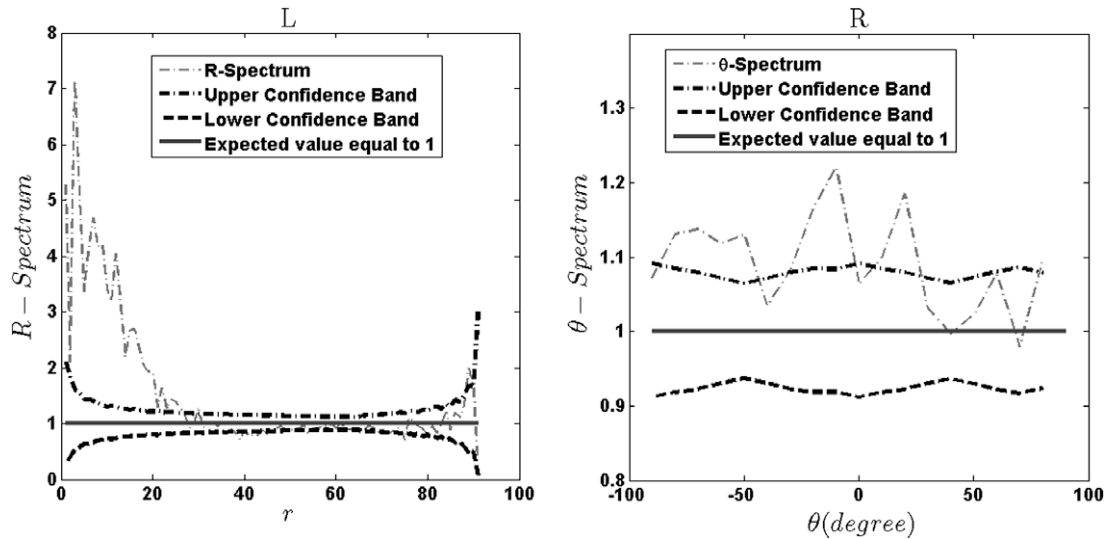


Fig. 8. 90% Confidence envelope of MTCP. (L)  $R$ -Spectrum, (R)  $\theta$ -Spectrum.

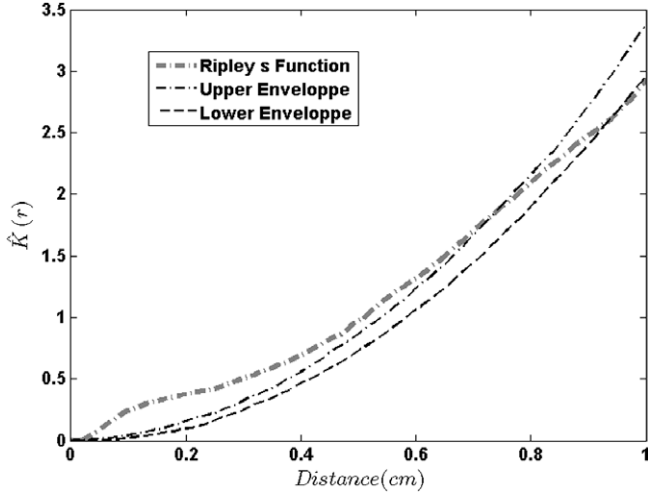


Fig. 9. 90% Confidence envelope of Ripley's method of MTCP.

we remark that the  $R$ -Spectrum reveals an aggregated structure for all the values  $r \leq 26$  before falling inside the confidence envelopes. This scale corresponds to  $4/26 = 0.154$  cm. Finally, the MTCP is isotropic. This property was not observed in Fig. 8(R) since the  $\theta$ -Spectrum does not fall inside the confidence band. Schabenberger et al. [14] point out this problem and propose, in order to interpret correctly the  $\theta$ -Spectrum, it is necessary to calculate the ratio  $\frac{f_{\theta}(\theta)}{f_{\theta}(-\theta)}$  which, under isotropy, follows the Fisher's distribution. This calculation is not presented here, since our main purpose in this work is to analyse spatial structure regardless of its orientation.

As mentioned previously, Ripley's method is performed and its result is shown in Fig. 9.

For aggregated structure, it is expected that Ripley's function falls above the upper confidence. As shown in Fig. 9, the aggregated structure is revealed (particularly for small distances) which confirms the result of spectral analysis. However, this method depends on the number of Monte-Carlo simulations. Thus, a small number of simulations can lead to false interpretations.

### 3.4. Regular structure

One of the methods to generate regular structure is to simulate simple sequential spatial inhibition denoted by SSI [18] for which each new point is generated uniformly and independently of all

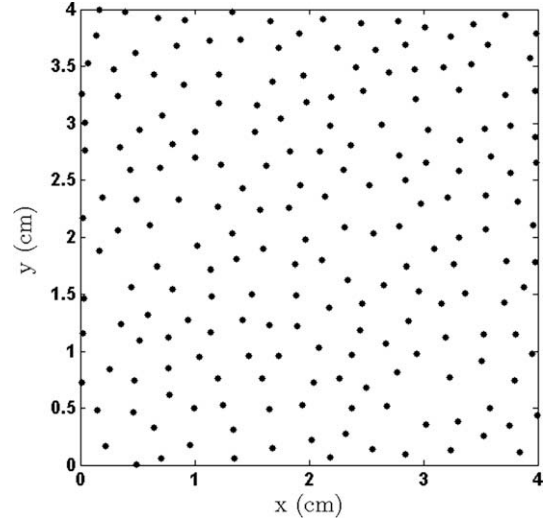


Fig. 10. Example of a SSI realization.

the previous points. News points are rejected as long as they lie closer than a fixed distance  $\delta$  from existing points. The procedure ends when no further points can be added in the studied region [13]. Fig. 10 shows 200 points generated under this process with  $\delta = 0.22$  cm. It should be noticed that this process is isotropic.

As expected, the generated points are regularly distributed within the square  $[0,4] \times [0,4]$ . The periodogram of this structure is shown in Fig. 11 and the spectrums are presented in Fig. 12.

For regular structures, the periodogram gives small values for small frequencies. That is clearly observed for the simulated pattern in Fig. 11. The  $R$ -Spectrum in Fig. 12(L) is under the lower confidence bound for small values of  $r$  which characterizes the regular structures. Furthermore, a peak is observed at  $r = 18$ . This value is close to the expected number  $4/\delta$ . Finally, the  $\theta$ -Spectrum remains inside the confidence envelope which confirms the isotropy of the generated structure. This regularity was detected via Ripley's method which is illustrated in Fig. 13 where it can be observed that the first non null value of Ripley's function occurs just beyond 0.22 cm which corresponds to the minimum permitted inter-event distance  $\delta$ .

More examples of application of the  $\theta$ -Spectrum are presented in the references [12,15]. The relevance of spectral analysis has been proved by using numerical simulations and the interpretations are consistent with the simulated structures. In addition to

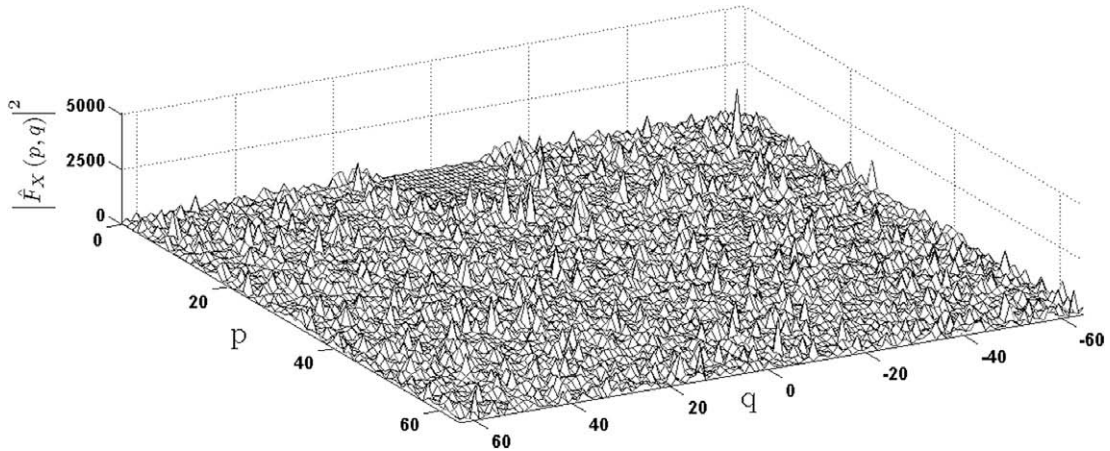


Fig. 11. Periodogram of SSI.



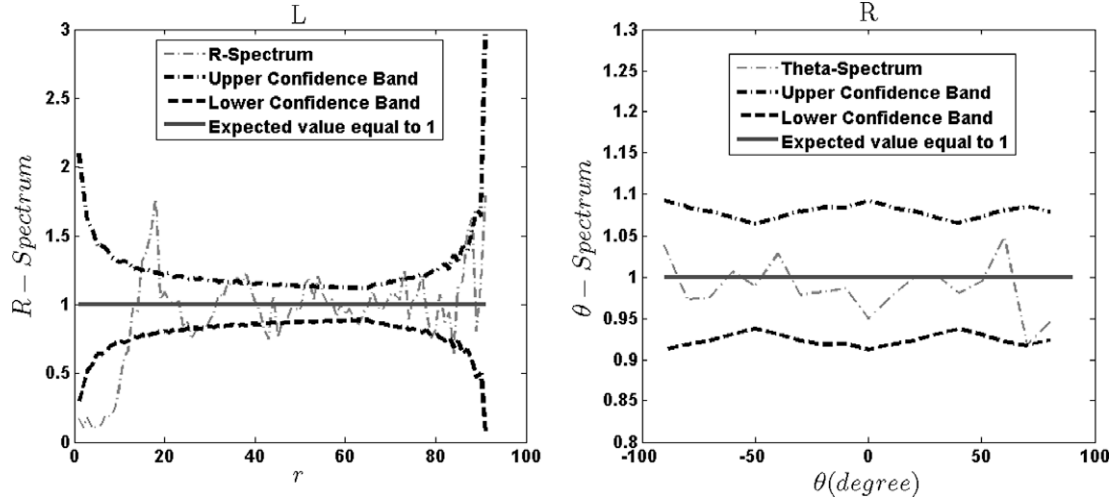


Fig. 12. 90% Confidence envelope of SSL. (L) R-Spectrum. (R)  $\theta$ -Spectrum.

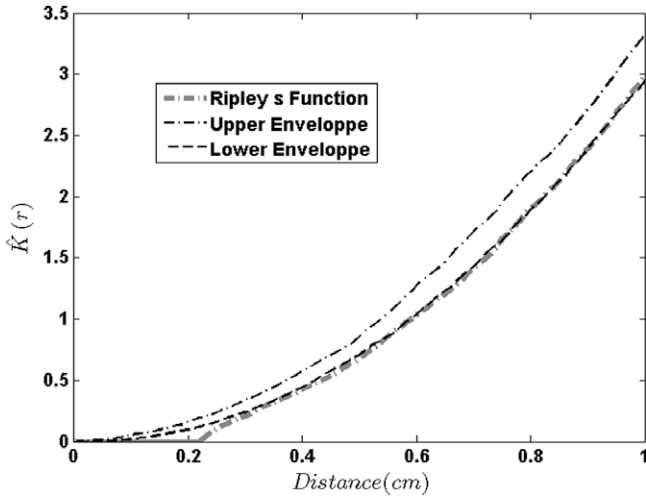


Fig. 13. 90% Confidence interval of Ripley's method of SSL.

distinguish between different types of patterns, spectral analysis can reveal more details not detected by classical methods. Furthermore, the orientation of patterns can be investigated.

#### 4. Application to experimental data and discussion

In this section we will discuss about the relevancy of CSR hypothesis as criterion to collect data. As previously presented, maximum pit depth is one the main design criterion in pitting corrosion and Gumbel's method is the most used to estimate its value for an objective area  $S$ . The method consists in subdividing an analyzed area to  $N$  cells with the same size  $s$  and to extract the maximum pit depths on each cell. These  $N$  values are used to estimate the maximum pit depth for a return period  $T$  corresponding to the objective area  $S$ . For a rigorous application of this method, the main theoretical assumptions are independence and homogeneity as explained previously. However, another condition difficult to satisfy and usually neglected is that every cell must contain the same number of pits. For this reason, an optimal choice of the size of cells is of prime importance. Many authors have discussed about this problem but no definitive method was proposed to determine this choice. Shibata [1] recommends that  $s$  should be chosen so as to contain plural number of pits, and notice that ex-

treme value distribution could not be applicable below a critical size of  $s$ .

Melchers [19] evokes the problems of "sampling representativeness" and notice that inhomogeneity of pit depths arises when metastable pitting is observed. He concludes that to collect data, large blocs are required for deeper depths to capture a sufficient number of extreme pits whereas small blocs when pits are small and numerous. A criterion to collect data was proposed by [20], consisting to choose samples for which the probability of metastable pitting occurrence not exceed 10%.

A CSR hypothesis of pits on a studied area assumes that there are no interactions between pits. Thus, the CSR structures can be considered as a model where the independence condition in terms of pit depths is satisfied. Therefore, it can be viewed as a relevant criterion to collect data since it is in agreement with independence hypothesis required in extreme value methods. Furthermore, when the analysed area is subdivided, same number of pits is expected in each cell because these pits are regularly distributed and the density in terms of number of pits per unit area is constant for each scale (subdivision). This property is made clear by simulating two patterns of 400 cm<sup>2</sup> where about 5000 pits coordinates are generated under CSR and MTCP process for which the parameters are chosen as  $\lambda_p = 550$ ,  $\mu = 10$  and  $\sigma = 0.6$  in order to have a final number of points close to 5000. These structures are subdivided into several cells which the numbers vary from 10 to 400. Afterwards, we test if a difference exists between the numbers of pits falling in each cell via the Chi-square test. The result of this simulation is shown in Fig. 14.

For all the subdivisions taken in this simulation, the CSR structure in Fig. 14(L), allows getting cells with the same number of pits (at given risk  $\alpha$ ). Whereas, with aggregated structure shown in Fig. 14(R), it is more difficult to get uniform subdivision in terms of number of pits. This remark leads to suppose that CSR structure as criterion to collect data is appropriate in order to respect the conditions of Gumbel's method. Testing the accuracy of this criterion and its effect in the estimation of the maximum pit depth based on numerical simulations will be presented in a next paper.

The CSR hypothesis was tested on the data collected from the 5 squares of 4 cm<sup>2</sup> related to the corroded aluminium sheet presented in Fig. 1. This experiment consists in immersing during 10 h, an aluminium sheet of area  $S$  equals to 150 cm<sup>2</sup> and thickness of 0.8 mm in a corrosive solution under ambient temperature and without agitation. The sheet was annealed during one hour at 300 °C, and the composition of the corrosive solution is: 0.5 g of NaCl, 4 g of FeSO<sub>4</sub>, 25 cm<sup>3</sup> of H<sub>2</sub>SO<sub>4</sub>, 200 cm<sup>3</sup> of H<sub>2</sub>O.

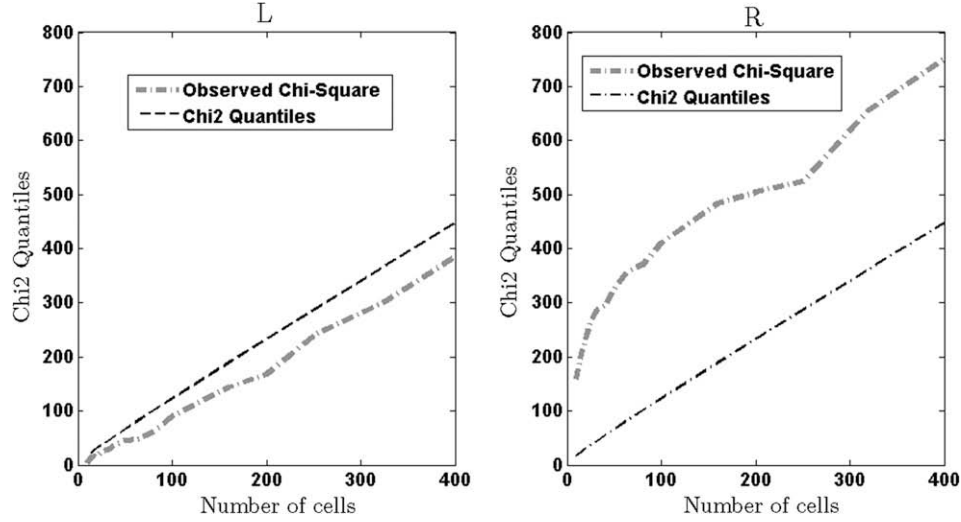


Fig. 14. 90% Chi-square test to compare the number of pits per cell. (L) CSR structure, (R) aggregated structure.

Fig. 15 shows the shape of three pits extracted from the aluminium sheet as like as their depths.

Only the  $R$ -Spectrum is calculated which is sufficient to accept or reject the CSR hypothesis. Thus, the 5 squares are considered as the analysed areas. They will be used to estimate the maximum pit depths in  $S = 150 \text{ cm}^2$ . It is worth noting that pit depths have been measured using an optical microscope with a  $20 \times$  magnification. The depth is calculated as the height separating the levels corresponding to the top (neat image) and the bottom (blurred image) of the pit. So, 248 pit depths are recorded which fall within the mask selected initially.

Fig. 16 shows the result of the  $R$ -Spectrum. We note that this  $R$ -Spectrum was performed for  $p_{\max} = q_{\max} = 16$  because of the smallest size of samples composed of the 5 squares (an average of 48 pits per squares).

Except some slight peaks beyond the confidence band, the pits within the squares  $S_1$ ,  $S_2$ ,  $S_4$  and  $S_5$  are randomly distributed and the CSR hypothesis cannot be rejected since the  $R$ -Spectrum re-

mains approximately inside the confidence bands. However, pits within  $S_3$  are aggregated since for  $r = 1$ , the  $R$ -Spectrum falls above the upper confidence band. This result may be explained by a high density of pit corrosion in  $S_3$  due, for example, to the local variation of the material composition in terms of density of inclusions leading to local variations of electrochemical conditions.

In order to apply Gumbel's method, the 5 squares must be subdivided into several cells where the pits are independent. Referring to the  $R$ -Spectrum in Fig. 16, we choose a subdivision for which the pits are distributed under CSR. For example at  $r = 2$ , the 5 squares verify this condition. This value of  $r$  corresponds to the scale  $2/2 = 1 \text{ cm}$ , then a subdivision in 4 cells allows to respect the conditions of Gumbel's method. This result is shown in Fig. 17.

In order to estimate the maximum pit depth for the return period  $T = 150$ , the parameters of Gumbel's distribution function were estimated by using the linear regression method. So, this estimation gives the value of  $298 \mu\text{m}$  with the 95% confidence interval  $[288 \mu\text{m}, 309 \mu\text{m}]$  which corresponds to the vertical line in Fig. 17.

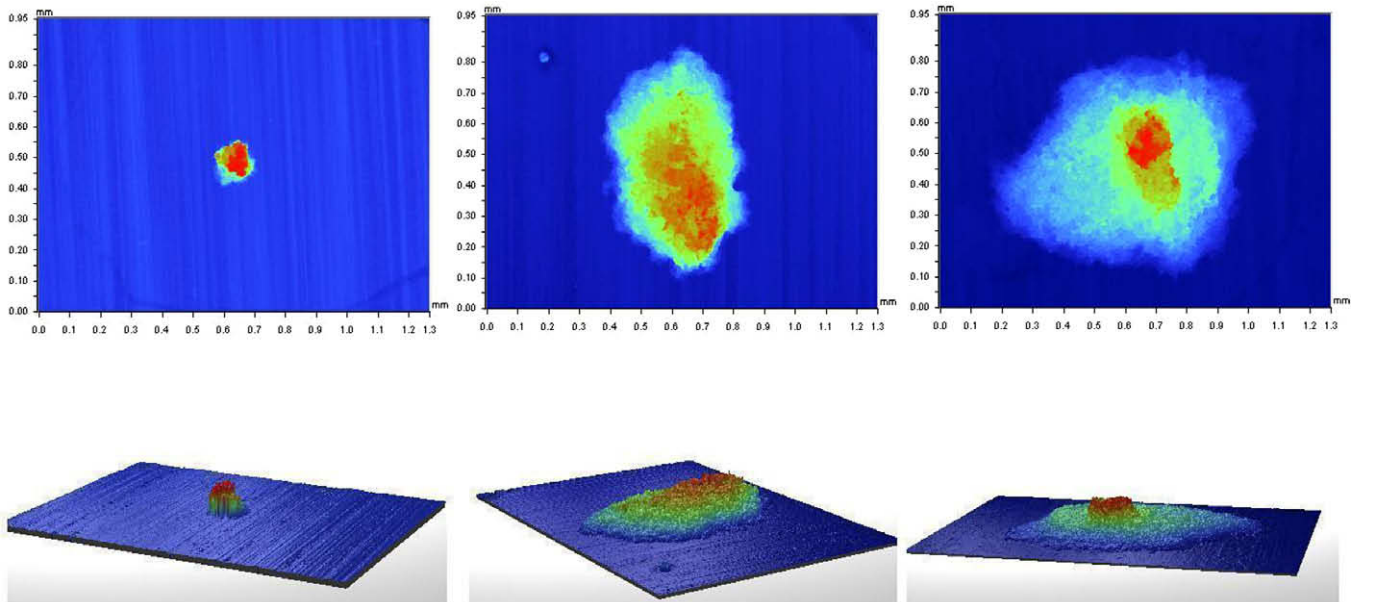


Fig. 15. Example of the structures of three pits extracted from the aluminium sheet.

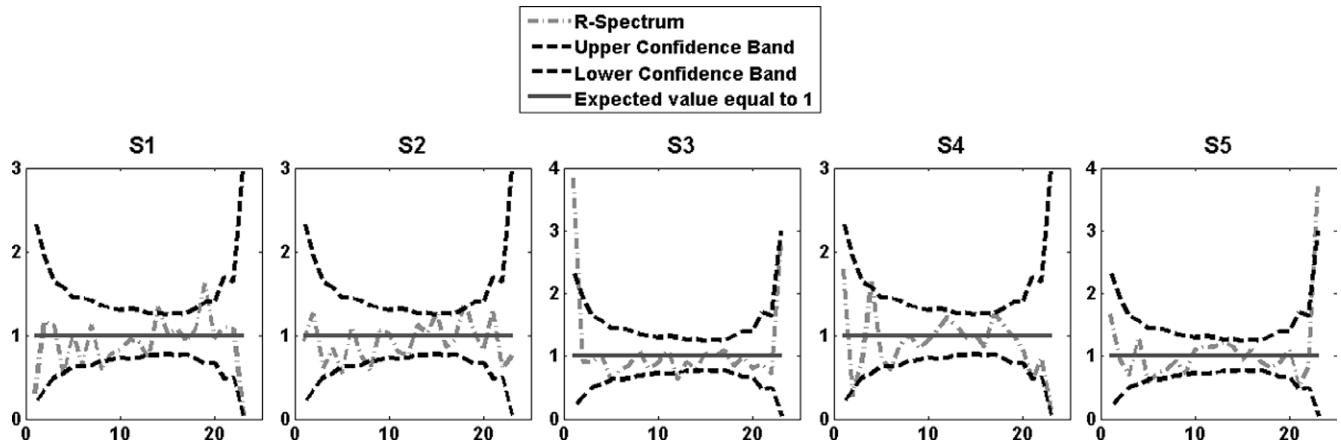


Fig. 16. R-Spectrum for the squares  $S_1$ ,  $S_2$ ,  $S_3$ ,  $S_4$  and  $S_5$ .

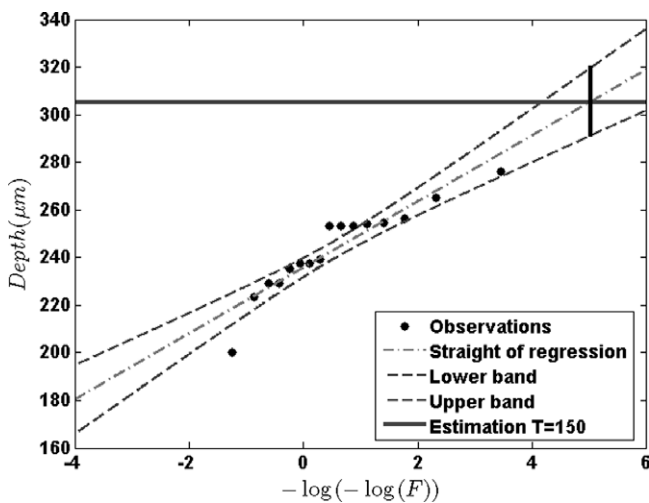


Fig. 17. Gumbel's method of the 5 squares, the vertical line corresponds to the 95% confidence interval of the estimation.

It should be emphasized that the proposed method gives statistical information about the spatial structure related to a corrosion process but not on its physical mechanism. It enables to detect eventual existing interactions between corrosion pits but not explain how and why these interactions take place. The description of these interactions is based only on the location of pits over the studied area. So, the model is two-dimensional and neglect the three-dimensional aspect of pitting corrosion. Such a simple two-dimensional model has been already considered in corrosion studies by several authors [2–4,10]. So, the spatial arrangement of pit structure leads to detect the presence of interactions between pits (for aggregated or regular structure) or supposes that pits are independent (for CSR structure). However, for a complete study of the phenomenon, marked process appear more appropriate because it take into account the characteristics of pits as depth, area or volume. In this case, the description is three-dimensional. This model will be presented in a next paper.

## 5. Conclusion

It is shown in this paper that spectral analysis is a powerful technique to distinguish between spatial data. The main advantage of this method is that it reveals more details about the structures without needing Monte-Carlo simulations compared with Ripley's

method. The periodogram gives a general comprehension about the structure but the calculation of the R-Spectrum and the  $\theta$ -Spectrum allow to extract more information particularly the scale for which the pattern is aggregated and the privileged orientation of pitting corrosion. The relevance of this method has been made clear for all the studied cases for which different spatial structures may be observed at different scales. Finally, we proposed that CSR criterion allows looking at the theoretical conditions of Gumbel's method in the aim of a rigorous estimation of the maximum pit depth.

More and more detection of pit interactions integrate pitting corrosion analysis since it reveals hidden information in the pattern. Nevertheless, the investigation of interactions via the statistical method presented in this paper does not allow to identify the source or the parameters leading to these interactions. The objective is only to accept or reject the dependence between pits which is useful when Gumbel's method is applied, and model cannot help us to understand the correlation between experimental conditions and the spatial structure of pits; it is only a statistical post-treatment in order to make clear the pit dependence or not at different scales of the corroded structure.

## Acknowledgements

A particular thanks to Professor Maxence Bigerelle (UTC, Compiègne, France) and to research and development engineer Dr. Benjamin Fournier (CEA, Saclay, France) for their interesting discussions and useful advices about the scientific problem of pitting corrosion.

## References

- [1] T. Shibata, Statistical and stochastic approaches to localised corrosion, *Corros. Sci.* 52 (1996) 813–830.
- [2] J. Lopez De La Cruz, R.H.A. Lindelauf, L. Koene, M.A. Gutiérrez, Stochastic approach to the spatial analysis of pitting corrosion and pit interaction, *Electrochem. Commun.* 9 (2007) 325–330.
- [3] N.D. Budiansky, L. Organ, J.L. Hudson, J.R. Scully, Detection of interactions among localised pitting sites on stainless steel using spatial statistics, *J. Electrochem. Soc.* 152 (2005) B152–B160.
- [4] N.R. Cawley, D.G. Harlow, Spatial statistics of particles and corrosion pits in 2024-T3 aluminium alloy, *J. Mater. Sci.* 31 (1996) 5127–5134.
- [5] M.S. Bartlett, The spectral analysis of two dimensional point processes, *Biometrika* 51 (1964) 299–311.
- [6] P. Couteron, Using spectral analysis to confront distribution of individual species with an overall periodic pattern in semi-arid vegetation, *Plant Ecol.* (2001) 1–15.
- [7] E. Renshaw, Spectral analysis in spatial analysis, *Forest Ecol. Manag.* 94 (1997) 165–174.
- [8] J. Illian, A. Penttinen, H. Stoyan, D. Stoyan, *Statistical Analysis and Modelling of Spatial Point Patterns*, John Wiley & Sons Ltd, 2008.

- [9] B.D. Ripley, Statistical Inference for Spatial Processes, Cambridge University Press, 1988.
- [10] J. Lopez De La Cruz, M.A. Gutiérrez, Spatial statistics of pitting corrosion patterning: quadrat counts and the non-homogeneous Poisson process, *Corros. Sci.* 50 (2008) 1441–1448.
- [11] D.R. Cox, Role of significance test, *Scand. J. Stat.* 4 (1977) 49–70.
- [12] M.A. Muggleston, E. Renshaw, A practical guide to the spectral analysis of spatial point processes, *Comput. Stat. & Data An.* 21 (1996) 43–65.
- [13] P.J. Diggle, Statistical Analysis of Spatial Point Patterns, first ed., Academic Press, London, 1983.
- [14] O. Schabenberger, C.A. Gotway, Statistical Methods for Spatial Data Analysis, CRC Press, US, 2005.
- [15] E. Renshaw, E.D. Ford, The description of spatial pattern using two-dimensional spectral analysis, *Vegetation* (1984) 75–85.
- [16] E. Renshaw, E.D. Ford, The interpretation of process from pattern using two-dimensional spectral analysis: methods and problems of interpretation, *Appl. Stat.* 32 (1983) 51–63.
- [17] M.A. Muggleston, E. Renshaw, Spectral tests of randomness for spatial point patterns, *Environ. Ecol. Stat.* 8 (2001) 237–251.
- [18] B.D. Ripley, Spatial Statistics, John Wiley & Sons, New York, 1981.
- [19] R.E. Melchers, Representation of uncertainty in maximum depth of marine corrosion pits, *Struct. Saf.* 27 (2005) 322–334.
- [20] R.E. Melchers, Extreme value statistics and long-term marine pitting corrosion of steel, *Probabilist. Eng. Mech.* 23 (2008) 482–488.